

## Bounds in coding

18<sup>th</sup> November 2005

**Definition 1.** A [prime power] is a prime or an integer power of a prime.

§

**Example 1.** Examples of prime powers are,

$$2, 3, 4, 5, 7, 8, 9, 11, 13, 16, 17, \dots$$

**Definition 2.** Let the alphabet be  $\mathbf{F}_q$ , in other words a Galois field  $GF(q)$ , where  $q$  is a prime power, and let the vector space  $V(n, q)$  be  $(\mathbf{F}_q)^n$ . Then a *linear code* over  $GF(q)$ , for some positive integer  $n$ , is a subspace of  $V(n, q)$ .

§

**Theorem 1.** A subset  $C$  of  $V(n, q)$  is a linear code if and only if,

- $\mathbf{u} + \mathbf{v} \in C$  for all  $\mathbf{u}$  and  $\mathbf{v}$  in  $C$
- $a\mathbf{u} \in C$  for all  $\mathbf{u} \in C$  and  $a \in GF(q)$

**Proof.** The proof follows from Definition 2 since, if  $C$  is a field, it must be closed under addition and multiplication. ¶

**Example 2.** A binary code is linear if and only if the sum of any two code words is a code word.

**Definition 3.** A *vector space*  $V$  is a set which is closed under finite vector addition and scalar multiplication. If the scalars are members of a field  $F$ , then  $V$  is called a vector space under  $F$ . Furthermore,  $V$  is a vector space under  $F$  if and only if for all members of  $V$  and  $F$  the following properties hold under addition,

- commutativity
- associativity
- existence of an identity
- existence of an inverse

while under multiplication the following,

- associativity under scalar multiplication
- distributivity of scalar sum
- distributivity of vector sum
- existence of a scalar multiplication identity

In other words, for all  $\mathbf{x}$ ,  $\mathbf{y}$  and  $\mathbf{z}$  in  $V$  and all  $p$  and  $q$  in  $F$ ,

- $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
- $\mathbf{0} + \mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{x}$
- $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$
- $r(s\mathbf{x}) = (rs)\mathbf{x}$
- $(r + s)\mathbf{x} = r\mathbf{x} + s\mathbf{x}$
- $r(\mathbf{x} + \mathbf{y}) = r\mathbf{x} + r\mathbf{y}$
- $1\mathbf{x} = \mathbf{x}$

§

**Example 3.** Let  $q$  be a prime power, and let  $GF(q)$  denote a finite field over  $q$  elements. Then, by *vector space over finite field* we mean a set  $GF(q)^n$  of all ordered  $n$ -tuples over  $GF(q)$ , which is closed under finite vector addition and multiplication, that is to say, multiplication by some scalar  $a$  in  $GF(q)$ .

**Theorem 2.** A non-empty subset  $C$  of  $V(n, q)$  is a subspace if and only if  $C$  is closed under addition and scalar multiplication. In other words

**Proof.** What Theorem 2 states amounts to saying that a non-empty  $C$  in  $V(n, q)$  is a subspace if and only if,

- $\mathbf{x}, \mathbf{y} \in C$  implies  $\mathbf{x} + \mathbf{y} \in C$
- if  $a \in GF(q)$  and  $\mathbf{x} \in C$ , then  $a\mathbf{x} \in C$

All properties to be met in Definition 3 are the same for  $C$  as for  $V(n, q)$  itself, provided that  $C$  is closed under addition and scalar multiplication. Therefore statements (a) and (b) are necessary for  $C$  to be a subspace. They are also sufficient since  $C$  is already a subset of  $V(n, q)$ . ¶

**Definition 4.** A linear combination of  $r$  vectors  $\mathbf{v}_1, \dots, \mathbf{v}_r$  in  $V(n, q)$  is any vector of the form  $\sum_{i=1}^r a_i \mathbf{v}_i$ , where  $a_i$  are scalars. Let  $A$  be a set of vectors  $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ . Then  $A$  is said to be *linearly dependent* if there exist scalars  $a_1, \dots, a_r$  not all of which are zero, such that  $\sum_{i=1}^r a_i \mathbf{v}_i = \mathbf{0}$ . And  $A$  is *linearly independent* if it is not linearly dependent, that is to say, if  $\sum_{i=1}^r a_i \mathbf{v}_i = \mathbf{0}$  implies  $a_i$  are all zero for  $i = 1, \dots, r$ .

§

**Definition 5.** Let  $C$  be a subspace of a vector space  $V(n, q)$  over  $GF(q)$ . Then a subset  $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$  of  $C$  is called a *generating-* or *spanning set* of  $C$  if every vector in  $C$  can be expressed as a linear combination of  $\mathbf{v}_1, \dots, \mathbf{v}_r$ . A *basis* of  $C$  is a generating set of the same which is also linearly independent.

§

**Definition 6.** For a  $q$ -ary  $(n, m, d)$ -code  $C$ , the *relative minimum distance* of  $C$  is defined to be

$$\delta(C) = \frac{d-1}{n}$$

§

**Definition 7.** Let a code alphabet  $A$  be of size  $q > 1$ ,  $n$  the size of each word,  $d$  the minimum distance, and  $A_q(n, d)$  the largest possible vocabulary size  $m$  such that there exists an  $(n, m, d)$ -code over  $A$ . Then any  $(n, m, d)$ -code  $C$  which has  $m = A_q(n, d)$  is called an *optimal code*. The *main coding theory problem* is precisely to find the value of  $A_q(n, d)$ .

§

**Definition 8.** Consider each word as an  $n$ -tuple. Then all such tuples lying within Hamming distance  $r$  of an  $n$ -tuple  $x$  are said to be within a *Hamming sphere* of radius  $r$  around  $x$ .

§

**Theorem 3.** Let the size of the alphabet be  $q = |A|$ , the size of a word be  $n$ , and the Hamming- or minimum distance be  $d$ . Then the Hamming- or sphere-packing bound on the size  $m$  of a code dictionary  $C$  is given by,

$$m \leq \frac{q^n}{\sum_{i=0}^{r'} (q-1)^i \binom{n}{i}}$$

where

$$r' = \left\lfloor \frac{d-1}{2} \right\rfloor$$

**Proof.** Let  $c$  be a code word. Let  $e(x, y)$  be the number of places which are different between two words  $x$  and  $y$ . Since there are  $q-1$  possibilities for each differing position between any two words, there are  $(q-1)^i$  possible errors when  $i$  places are different. And to position these  $i$  places there are altogether  $\binom{n}{i}$  ways. Therefore the number of all words  $w_i$  such that  $e(w_i, c) \leq r$  is the number  $n_r$  of  $n$ -tuples in a Hamming sphere of radius  $r$  around  $c$ , and is,

$$n_r = \sum_{i=0}^r (q-1)^i \binom{n}{i}$$

Then the lower bound for our code is  $d(C) > 2r$ , that is to say,  $d(C) \geq 2r + 1$ . In other words, Hamming spheres of radius  $r$  around the  $m$  code words of  $C$  are mutually nonintersecting. There are a total of  $q^n$  possible  $n$ -tuples, that is words of length  $n$ , not all of which are code words. In other words,  $m < q^n$ . And since there are  $n_r$  of these  $n$ -tuples within each sphere, the the number of the all the  $n$ -tuples contained within the space of all these  $n$ -tuples over the alphabet  $A$  is  $n_r m$ . Hence,

$$m \sum_{i=0}^r (q-1)^i \binom{n}{i} \leq q^n$$

and thus this theorem. ¶

**Definition 9.** Codes which satisfies the Hamming bound are called [perfect codes].

§

**Problem 1.** Let  $r$  and  $n$  be integers such that  $0 < r < \frac{n}{2}$ , then prove that,

$$\left[ 8n \left( \frac{r}{n} \right) \left( 1 - \frac{r}{n} \right) \right]^{-\frac{1}{2}} 2^{nH\left(\frac{r}{n}, 1 - \frac{r}{n}\right)} \leq \sum_{i=0}^r \binom{n}{i} \leq 2^{nH\left(\frac{r}{n}, 1 - \frac{r}{n}\right)}$$

where  $H(x, y)$  is the entropy function the arguments  $x$  and  $y$  of which are probabilities and  $H(\cdot, \cdot)$  has the unit of bits per symbol. (Hint: Stirling's approximation to  $n!$ , cf MacWilliams and Sloane, 1977)

§

**Note 1.** Let  $C(n, d)$  be a code with words of length  $n$  and minimum distance between words  $d$ . Let  $m_{n,d}$  be the number of code words in  $C(n, d)$ . Then the size of the largest dictionary of  $n$ -tuples with fractional minimum distance  $d_f$  is,

$$m_m(n, d_f) = \max_{\{C(n, d): (\frac{d}{n}) \geq d_f\}} |C(n, d)|$$

§

**Problem 2.** From Note 1, show that for  $n$  fixed,  $m_m(n, d_f)$  is a monotonous nonincreasing function of  $d_f$ . Then show that with  $d_f$  fixed,  $m_m(n, d_f)$  increases exponentially with  $n$ .

§

**Definition 10.** The *asymptotic transmission rate* is defined to be,

$$R(d_f) \lim_{n \rightarrow \infty} \frac{1}{n} \log m_m(n, d_f)$$

Also defined are the upper- and the lower bounds on this rate,

$$\bar{R}(d_f) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log m_m(n, d_f)$$

and

$$\underline{R}(d_f) = \liminf_{n \rightarrow \infty} \frac{1}{n} \log m_m(n, d_f)$$

§

**Note 2.** For large  $n$ , show that  $\underline{R}(d_f) < R(d_f) < \bar{R}(d_f)$ .

§

**Example 4.** Using the results from Problem 1 we obtain the Hamming bound for the binary code,

$$m \leq \left( 8n \left( \frac{r}{n} \right) \left( 1 - \frac{r}{n} \right) \right)^{\frac{1}{2}} 2^{n(1-H(\frac{r}{n}, 1-\frac{r}{n}))} \quad (1)$$

where

$$r = \left\lfloor \frac{d-1}{2} \right\rfloor$$

Equation 1 must hold for all binary dictionaries, therefore it gives an upper bound on the maximum dictionary size  $m_m(n, d_f)$  over all dictionaries whose word length is  $n$  and fractional distance,

$$d_f = \frac{d}{n} = \frac{2r + \left\{ \frac{1}{2} \right\}}{n}$$

where the choice of 1 or 2 depends on whether  $d$  is odd or respectively even. For large  $n$ ,

$$m_m(n, d_f) \leq \left( 9n \left( \frac{d_f}{2} \right) \left( 1 - \frac{d_f}{2} \right) \right)^{\frac{1}{2}} 2^{n \left( 1 - H\left(\frac{d_f}{2}, 1 - \frac{d_f}{2}\right) \right)}$$

The upper bound for the attainable information rate is,

$$\begin{aligned}\bar{R}(d_f) &= \limsup_{n \rightarrow \infty} \frac{1}{n} \log_2 m_m(n, d_f) \\ &\leq \lim_{n \rightarrow \infty} \left\{ \frac{1}{2} \frac{\log_2 n}{n} + \frac{1}{2n} \log_2 \left( \frac{9d_f}{2} \left( 1 - \frac{d_f}{2} \right) \right) \right\} + 1 - H \left( \frac{d_f}{2}, 1 - \frac{d_f}{2} \right)\end{aligned}$$

As  $n$  approaches infinity,

$$\bar{R}(d_f) \leq 1 - H \left( \frac{d_f}{2}, 1 - \frac{d_f}{2} \right)$$

**Problem 3.** Work out the details of derivation of Example 4.

§

**Theorem 4.** Let  $d(c_i, c_j)$  be the Hamming distance between the code words  $c_i$  and  $c_j$ . Let  $d(C)$  be the minimum distance between code words, and  $\bar{d}$  the average distance between words. If,

$$\frac{d}{n} > \frac{q-1}{q}$$

then the *Plotkin's bound*,

$$m_{n,d} \leq \frac{\frac{d}{n}}{\frac{d}{n} - \frac{q-1}{q}}$$

**Proof.** The average distance gives an upper bound for the minimum distance, that is  $d \leq \bar{d}$ , where

$$\begin{aligned}\bar{d} &= \frac{\sum_{i=2}^m \sum_{j=1}^{i-1} d(c_i, c_j)}{\sum_{i=2}^m \sum_{j=1}^{i-1} 1} \\ &= \left( \frac{m(m-1)}{2} \right)^{-1} \sum_{i=2}^m \sum_{j=1}^{i-1} d(c_i, c_j)\end{aligned}$$

Since the Plotkin's bound is an upper bound on  $d$ , we need to maximise,

$$\begin{aligned}\sum_{i>j} d(c_i, c_j) &= \sum_{i>j} \sum_{k=1}^n d(c_{ik}, c_{jk}) \\ &= \sum_{k=1}^n \sum_{i>j} d(c_{ik}, c_{jk})\end{aligned}$$

This implies (*cf* Plotkin, 1960),

$$\sum_{i>j} d(c_i, c_j) \leq \sum_{k=1}^n \max_{\{c_{ik}, i=1, \dots, m\}} \left\{ \sum_{i>j} d(c_{ik}, c_{jk}) \right\}$$

which says that the upper bound is maximised by choosing a maximising  $c_{ik}$  from the alphabet  $A$ . However this is,

$$\max_{c_{ik}, i=1, \dots, m} \sum_{i>j} d(c_{ik}, c_{jk}) \leq \left( \frac{m}{n} \right)^2 \frac{n(n-1)}{2}$$

Providing that,

$$\frac{d}{n} > \frac{n-1}{n}$$

then

$$d \leq n \left( \frac{m}{m-1} \right) \left( \frac{n-1}{n} \right)$$

¶

**Note 3.** Notice how,

$$\sum_{i=1}^{m-1} \sum_{j=i+1}^m (\cdot) = \sum_{i=2}^m \sum_{j=1}^{i-1} (\cdot)$$

Equivalently to this are,

$$\sum_{i < j} (\cdot) \quad \text{and} \quad \sum_{i > j} (\cdot)$$

§

**Problem 4.** Prove Note 3 on double summations.

§

**Note 4.** If,

$$d_f > \frac{q-1}{q}$$

then,

$$m_m(n, d_f) \leq \frac{d_f}{d_f - \left(\frac{q-1}{q}\right)}$$

and then,

$$\bar{R}(d_f) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log m_m(n, d_f) = 0$$

On the other hand if,

$$d_f \leq \frac{q-1}{q}$$

then from,

$$m(n, d) = \sum_{a \in A} m_a(n, d)$$

where  $m(n, d) = |C(n, d)|$ ,  $C(n, d)$  being any code consisting of  $n$ -tuples whose minimum distance is at least  $d$ , and  $m_x(n, d) = |C_x(n, d)|$ ,  $C_x(n, d)$  comprising all  $n$ -tuples in  $C(n, d)$  which begin with the symbol  $x$ . Hence,

$$\begin{aligned} m(n, d) &\leq qm_x(n, d) \\ &= qm(n-1, d) \\ &\vdots \\ &= q^{n-k}m(k, d) \end{aligned}$$

Provided  $k$  is small enough, we may yet use the Plotkin's bound, hence

$$m(n, d) \leq \frac{q^{n-k} \binom{d}{k}}{\binom{d}{k} - \left(\frac{q-1}{q}\right)}$$

when

$$\frac{d}{k} > \frac{q-1}{q}$$

Choose  $k$  the largest integer satisfying

$$\frac{d}{k} - \frac{1}{qk} \geq \frac{q-1}{q}$$

Then,

$$k + r = \frac{qd - 1}{q - 1}$$

where  $0 \leq r < 1$ . And then,

$$\frac{m(n, d) \leq q^{n - \left(\frac{qd-1}{q-1}\right)} r + 1d}{(q-1)r + 1}$$

Finally,

$$m(n, d) \leq q^{n - \left(\frac{qd-1}{q-1}\right)d}$$

and, if  $d_f$  is fixed and  $n$  become large,

$$\bar{R}(d_f) \leq \log q \left(1 - \frac{q}{q-1}d_f\right)$$

§

**Problem 5.** Prove that,

$$q^r ((q-1)r + 1)^{-1} \leq 1$$

for  $0 \leq r \leq 1$

§

**Proposition 1.** Let  $C$  be a code containing binary  $n$ -tuples,  $m_d(x)$  the number of code words within distance  $d$  of an  $n$ -tuple  $x$ . Further, let  $A$  be a new code whose code words are the difference vectors  $a_1, \dots, a_{m_d}$  such that  $a_i = c_i \ominus x$ ,  $i = 1, \dots, m_d$ , where  $\ominus$  denotes modulo subtraction of the vectors, element by element. Assume that  $d < \frac{n}{2}$  and both  $d$  and  $m$  are large enough such that  $m_d(x) \geq 2$ . Then,

$$\frac{d_c}{n} \leq \frac{2d}{n} \left(1 - \frac{d}{n}\right) \frac{m_a}{m_a - 1} \quad (2)$$

where

$$m_a \geq \left\lceil 2^{-n} m \sum_{i=0}^d \binom{n}{i} \right\rceil$$

**Proof.** Since  $C$  is a code of binary  $n$ -tuples, there are

$$\sum_{i=0}^d \binom{n}{i}$$

$n$ -tuples within distance  $d$  of each code word. This gives the total of

$$m \sum_{i=0}^d \binom{n}{i}$$

$n$ -tuples in the Hamming sphere around the  $m$  code words.

There are  $m_d(x)$  code words within the distance  $d$  of any  $n$ -tuple  $x$ . For  $x$  in  $X^n$ ,  $c$  in  $C$  and  $d(x, c) \leq d$ , the number of pairs  $(x, c)$  can be counted by picking up first  $x$  and then  $c$ , hence

$$\sum_{x \in X^n} m_d(x) = m \sum_{i=0}^d \binom{n}{i}$$

Since  $X^n$  contains  $2^n$  of  $n$ -tuples, consequently there exists some value of  $x$  such that,

$$m_d(x) \geq \left\lceil 2^{-n} m \sum_{i=0}^d \binom{n}{i} \right\rceil$$

Let  $c_1, \dots, c_{m_d}$  be code words in  $C$  that lie within Hamming distance  $d$  of the  $n$ -tuple  $x$ . Consider the difference vector  $a_1, \dots, a_{m_d}$  such that  $a_i = c_i \ominus x$ . Then  $A$  is a set of localised code words of  $C$ . Then,

$$a_i \ominus a_j = (c_i \ominus x) \ominus (c_j \ominus x) = c_i \ominus c_j$$

and we have,

$$d(c_i, c_j) = d(a_i, a_j)$$

Thus,

$$m_a \geq m_d(x) \geq \left\lceil 2^{-n} m \sum_{i=0}^d \binom{n}{i} \right\rceil$$

Also,  $d_a \geq d_c$  and  $w(a_i) \leq d$  for all  $n$ -tuple  $a_i$  in  $A$ , where the Hamming weight  $w(a_i)$  is the number of nonzero elements in  $a_i$ .

Next, applying the average-distance Plotkin bound to the localised code  $A$  one obtains,

$$d_c \leq d_a \leq \bar{d}_a = \left( \frac{m_a(m_a - 1)}{2} \right)^{-1} \sum_{i>j} \sum d(a_i, a_j) \quad (3)$$

We maximise RHS of Equation 3 to get rid of the dependence on  $A$ . We enlarge our restriction on  $w(a_i)$  above to the set of all possible  $a_i$  in  $A$ , thus,

$$\sum_{a_i \in A} w(a_i) \leq m_a d \quad (4)$$

Then, let  $z_k$  be the number of code words in  $A$  having a 0 in the  $k^{\text{th}}$  position. We maximise,

$$\sum_{i>j} \sum d(a_i, a_j) = \sum_{k=1}^n (m_a - z_k) \quad (5)$$

subject to the constraint of Equation 4 that,

$$\sum_{k=1}^n (m_a - z_k) \leq m_a d \quad (6)$$

By setting,

$$z_k = \frac{m_a d}{n} \quad (7)$$

we maximise RHS of Equation 5 under the constraint in Equation 6. From Equation's 3, 5 and 7 we obtain Equation 2.  $\P$

**Algorithm 1** *Gilbert bound, a lower bound to  $m$  for  $n$ ,  $d$  and  $q$ .*

```

 $S^n \leftarrow X^n$ 
for all  $c_i$  in  $S^n$  do
  for all  $n$ -tuples  $c_j$  within  $d - 1$  distance of  $C$  do
    remove  $c_j$ 
  endfor
endfor

```

**Note 5.** For the Gilbert bound algorithm, Algorithm 1, initially  $|S|^n = |X|^n$ . For each  $c_i$  chosen, at most

$$\sum_{i=0}^{d-1} (q-1)^i \binom{n}{i}$$

$n$ -tuples are removed. If

$$(m-1) \sum_{i=0}^{d-1} (q-1)^i \binom{n}{i} < q^n$$

then the algorithm will not stop after  $m - 1$  code-word selections.

§

## Bibliography

- Solomon W Golomb, Robert E Peile and Robert A Scholtz. *Basic concepts in information theory and coding, the adventures of secret agent 00111*. Plenum Press, 1994
- Raymond Hill. *A first course in coding theory*. Clarendon, 1986
- San Ling and Chaoping Xing. *Coding theory, a first course*. Cambridge University Press, 2004
- M Plotkin. Binary codes with specified minimum distance. *IRE Transaction on Information Theory*. **6**, 445–50, 1960
- F J MacWilliams and N J A Sloane. *The theory of error-correcting codes*. North-Holland, Amsterdam, 1977